

1131-0420



#2 OIPE

RAW SEQUENCE LISTING

DATE: 02/14/2002

PATENT APPLICATION: US/10/010,408

TIME: 19:09:10

Input Set : N:\Crf3\RULE60\10010408.txt

Output Set: N:\CRF3\02142002\J010408.raw

SEQUENCE LISTING

4 (1) GENERAL INFORMATION:

6 (i) APPLICANT: John J. Castellot, Jr.

8 (ii) TITLE OF INVENTION: Novel Heparin-Induced CCN-Like Molecules
9 and Uses Therefor

11 (iii) NUMBER OF SEQUENCES: 13

13 (iv) CORRESPONDENCE ADDRESS:

14 (A) ADDRESSEE: LAHIVE & COCKFIELD, LLP

15 (B) STREET: 28 State Street

16 (C) CITY: Boston

17 (D) STATE: Massachusetts

18 (E) COUNTRY: USA

19 (F) ZIP: 02109

21 (v) COMPUTER READABLE FORM:

22 (A) MEDIUM TYPE: Floppy disk

23 (B) COMPUTER: IBM PC compatible

24 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

25 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

27 (vi) CURRENT APPLICATION DATA:

C--> 28 (A) APPLICATION NUMBER: US/10/010,408

C--> 29 (B) FILING DATE: 07-Dec-2001

35 (C) CLASSIFICATION:

38 (vii) PRIOR APPLICATION DATA:

33 (A) APPLICATION NUMBER: 09/044,273

34 (B) FILING DATE: March 19, 1998

39 (A) APPLICATION NUMBER:

40 (B) FILING DATE:

42 (viii) ATTORNEY/AGENT INFORMATION:

43 (A) NAME: Amy E. Mandragouras

44 (B) REGISTRATION NUMBER: 36,207

45 (C) REFERENCE/DOCKET NUMBER: MBI-004

47 (ix) TELECOMMUNICATION INFORMATION:

48 (A) TELEPHONE: (617)227-7400

49 (B) TELEFAX: (617)742-4214

52 (2) INFORMATION FOR SEQ ID NO: 1:

54 (i) SEQUENCE CHARACTERISTICS:

55 (A) LENGTH: 1708 base pairs

56 (B) TYPE: nucleic acid

57 (C) STRANDEDNESS: single

58 (D) TOPOLOGY: linear

60 (ii) MOLECULE TYPE: cDNA

63 (ix) FEATURE:

64 (A) NAME/KEY: CDS

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/010,408

DATE: 02/14/2002

TIME: 19:09:10

Input Set : N:\Crif3\RULE60\10010408.txt

Output Set: N:\CRF3\02142002\J010408.raw

```

65      (B) LOCATION: 249..1001
68      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
70 GACGCTTCTG ATCTCCAGAG GACCCTGGGG TGGGACAGGG GCCTTGGCAA GGCTGCAGCC      60
72 GCTGGGCAGT GGCTTGAAT GGAGGTCTTT ATTACTGGGA ACTGAGGAGC TAAGAGGCTC      120
74 CTGTCAGCTT GTCCTAAAGT CTTAGCACTT GTGGTGGCTT GGGCTTCACA CACTGTCAGA      180
76 CACCTTCGTG GTGGCCTCCA CGGCCTCACC TTCAGGTTTG AAGCTGGCTC CACAAGGGAC      240
78 ACGGTGAC ATG AGG GGC AGC CCA CTG ATC CAT CTT CTG GCC ACT TCC TTC      290
79      Met Arg Gly Ser Pro Leu Ile His Leu Leu Ala Thr Ser Phe
80      1          5          10
82 CTC TGC CTT CTC TCA ATG GTG TGT GCC CAG CTG TGC CGG ACA CCC TGT      338
83 Leu Cys Leu Leu Ser Met Val Cys Ala Gln Leu Cys Arg Thr Pro Cys
84 15          20          25          30
86 ACC TGT CCT TGG ACA CCA CCC CAG TGC CCA CAG GGG GTA CCC CTG GTG      386
87 Thr Cys Pro Trp Thr Pro Pro Gln Cys Pro Gln Gly Val Pro Leu Val
88          35          40          45
90 CTG GAT GGC TGT GGC TGC TGT AAA GTG TGT GCA CGG AGG CTG GGG GAG      434
91 Leu Asp Gly Cys Gly Cys Cys Lys Val Cys Ala Arg Arg Leu Gly Glu
92          50          55          60
94 TCC TGC GAC CAC CTG CAT GTC TGC GAC CCC AGC CAG GGC CTG GTT TGT      482
95 Ser Cys Asp His Leu His Val Cys Asp Pro Ser Gln Gly Leu Val Cys
96          65          70          75
98 CAG CCT GGG GCA GGC CCT GGC GGC CAT GGG GCT GTG TGT CTC TTG GAT      530
99 Gln Pro Gly Ala Gly Pro Gly Gly His Gly Ala Val Cys Leu Leu Asp
100      80          85          90
102 GAG GAT GAC GGT AGC TGT GAG GTG AAT GGC CGC AGG TAC CTG GAT GGA      578
103 Glu Asp Asp Gly Ser Cys Glu Val Asn Gly Arg Arg Tyr Leu Asp Gly
104 95          100          105          110
106 GAG ACC TTT AAA CCC AAT TGC AGG GTC CTG TGC CGC TGT GAT GAC GGT      626
107 Glu Thr Phe Lys Pro Asn Cys Arg Val Leu Cys Arg Cys Asp Asp Gly
108          115          120          125
110 GGC TTC ACC TGC CTG CCG CTG TGC AGT GAG GAT GTG CGG CTG CCC AGC      674
111 Gly Phe Thr Cys Leu Pro Leu Cys Ser Glu Asp Val Arg Leu Pro Ser
112          130          135          140
114 TGG GAC TGC CCA CGC CCC AAG AGA ATA CAG GTG CCA GGA AAG TGC TGC      722
115 Trp Asp Cys Pro Arg Pro Lys Arg Ile Gln Val Pro Gly Lys Cys Cys
116          145          150          155
118 CCC GAG TGG GTA TGT GAC CAG GGA GTG ACA CCG GCG ATC CAG CGC TCC      770
119 Pro Glu Trp Val Cys Asp Gln Gly Val Thr Pro Ala Ile Gln Arg Ser
120          160          165          170
122 ACG GCG CAA GGA CAC CAA CTT TCT GCC CTT GTC ACT CCT GCC TCT GCT      818
123 Thr Ala Gln Gly His Gln Leu Ser Ala Leu Val Thr Pro Ala Ser Ala
124 175          180          185          190
126 GAT GCT CCT TGT CCA AAT TGG AGC ACA GCC TGG GGC CCC TGC TCA ACC      866
127 Asp Ala Pro Cys Pro Asn Trp Ser Thr Ala Trp Gly Pro Cys Ser Thr
128          195          200          205
130 ACC TGT GGG CTG GGC ATA GCC ACC CGA GTG TCC AAC CAG AAC CGA TTC      914
131 Thr Cys Gly Leu Gly Ile Ala Thr Arg Val Ser Asn Gln Asn Arg Phe
132          210          215          220
134 TGC CAA CTG GAG ATC CAA CGC CGC CTG TGT CTG CCC AGA CCC TGC CTG      962

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/010,408

DATE: 02/14/2002

TIME: 19:09:10

Input Set : N:\Crif3\RULE60\10010408.txt

Output Set: N:\CRF3\02142002\J010408.raw

```

135 Cys Gln Leu Glu Ile Gln Arg Arg Leu Cys Leu Pro Arg Pro Cys Leu
136          225          230          235
138 GCA GCC AGG AGC CAC AGC TCA TGG AAC AGT GCT TTC TAAGGCCAAC      1008
139 Ala Ala Arg Ser His Ser Ser Trp Asn Ser Ala Phe
140          240          245          250
142 TGGGGATGCG GATACAGGGC CTGCCATCCT CAGCAAATGA CCCTAGGACC AGGCCCTGGA      1068
144 CTGCTGGTAG ATGCTCTTCT CCATGCTCTT GGCTGCAGTT AACTGTCCTG CTTGGATTCA      1128
146 CTGTGTAGAG CCACTGAGCG ATCCCTGCTC TGTCTGAGGT AGGCGGAGCA GGTGACCAGC      1188
148 TCCAGTTCTC TGGTTCAGCC TGGAAATCTG GGTTCCTCCTG GCTCATTCCT CAAAACATCC      1248
150 CTGTACAAAA AGGACAACCA AAAAGACCTT TAAACCTAGG CTATACTGGG CAAACCTGGC      1308
152 CACCGTGCTG GGGATAAGGT CAATGTTAGG ACCAGACAGC AGATTGCCTG AAACCTCCAA      1368
154 TTCCCTTCTT GGAATTCTGT ATGCTTGTC CCAAAGATGA TGAATGAACT CGTAAGTGTA      1428
156 CCTTCCCTGA CCTGAGAACA CCCTGCCTGC TCGGGAAGTA TTCAGGGGCA GAATTCTCTG      1488
158 TGAACATGAA GAGATGAATC AACTGTCCT TAAGAAATTC CTGAAAGTCC AGGAACTTGA      1548
160 GCTTTGTATT TTCAGGAATG CACATCTCTT AAGCACTCGC AAAACAGGAA GGCTCCACAC      1608
162 CTCTGGCAGG CCAGGGCCTT TCTCTTCAGC ATGAGAAAGA CAAGGGACAG CAGAGTACTC      1668
164 TCCTCTGGAG GACTAGTCTA GCCTAGAATA AACACCCAAA      1708
167 (2) INFORMATION FOR SEQ ID NO: 2:
169     (i) SEQUENCE CHARACTERISTICS:
170         (A) LENGTH: 250 amino acids
171         (B) TYPE: amino acid
172         (D) TOPOLOGY: linear
174     (ii) MOLECULE TYPE: protein
176     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
178 Met Arg Gly Ser Pro Leu Ile His Leu Leu Ala Thr Ser Phe Leu Cys
179   1          5          10          15
181 Leu Leu Ser Met Val Cys Ala Gln Leu Cys Arg Thr Pro Cys Thr Cys
182          20          25          30
184 Pro Trp Thr Pro Pro Gln Cys Pro Gln Gly Val Pro Leu Val Leu Asp
185          35          40          45
187 Gly Cys Gly Cys Cys Lys Val Cys Ala Arg Arg Leu Gly Glu Ser Cys
188          50          55          60
190 Asp His Leu His Val Cys Asp Pro Ser Gln Gly Leu Val Cys Gln Pro
191   65          70          75          80
193 Gly Ala Gly Pro Gly Gly His Gly Ala Val Cys Leu Leu Asp Glu Asp
194          85          90          95
196 Asp Gly Ser Cys Glu Val Asn Gly Arg Arg Tyr Leu Asp Gly Glu Thr
197          100         105         110
199 Phe Lys Pro Asn Cys Arg Val Leu Cys Arg Cys Asp Asp Gly Gly Phe
200          115         120         125
202 Thr Cys Leu Pro Leu Cys Ser Glu Asp Val Arg Leu Pro Ser Trp Asp
203          130         135         140
205 Cys Pro Arg Pro Lys Arg Ile Gln Val Pro Gly Lys Cys Cys Pro Glu
206 145          150         155         160
208 Trp Val Cys Asp Gln Gly Val Thr Pro Ala Ile Gln Arg Ser Thr Ala
209          165         170         175
211 Gln Gly His Gln Leu Ser Ala Leu Val Thr Pro Ala Ser Ala Asp Ala
212          180         185         190
214 Pro Cys Pro Asn Trp Ser Thr Ala Trp Gly Pro Cys Ser Thr Thr Cys

```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/010,408

DATE: 02/14/2002
TIME: 19:09:10

Input Set : N:\Crif3\RULE60\10010408.txt
Output Set: N:\CRF3\02142002\J010408.raw

215		195		200		205	
217	Gly	Leu	Gly	Ile	Ala	Thr	Arg
218		210		215		220	
220	Leu	Glu	Ile	Gln	Arg	Arg	Leu
221	225			230		235	240
223	Arg	Ser	His	Ser	Ser	Trp	Asn
224				245		250	
225	(2) INFORMATION FOR SEQ ID NO: 3:						
227	(i) SEQUENCE CHARACTERISTICS:						
228	(A) LENGTH: 753 base pairs						
229	(B) TYPE: nucleic acid						
230	(C) STRANDEDNESS: single						
231	(D) TOPOLOGY: linear						
233	(ii) MOLECULE TYPE: cDNA						
236	(ix) FEATURE:						
237	(A) NAME/KEY: CDS						
238	(B) LOCATION: 1..750						
241	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:						
243	ATG	AGG	GGC	AGC	CCA	CTG	ATC
244	Met	Arg	Gly	Ser	Pro	Leu	Ile
245	1			5		10	15
247	CTT	CTC	TCA	ATG	GTG	TGT	GCC
248	Leu	Leu	Ser	Met	Val	Cys	Ala
249			20			25	30
251	CCT	TGG	ACA	CCA	CCC	CAG	TGC
252	Pro	Trp	Thr	Pro	Pro	Gln	Cys
253			35			40	45
255	GGC	TGT	GGC	TGC	TGT	AAA	GTG
256	Gly	Cys	Gly	Cys	Cys	Lys	Val
257		50				55	60
259	GAC	CAC	CTG	CAT	GTC	TGC	GAC
260	Asp	His	Leu	His	Val	Cys	Asp
261	65				70		75
263	GGG	GCA	GGC	CCT	GGC	GGC	CAT
264	Gly	Ala	Gly	Pro	Gly	Gly	His
265				85		90	95
267	GAC	GGT	AGC	TGT	GAG	GTG	AAT
268	Asp	Gly	Ser	Cys	Glu	Val	Asn
269			100			105	110
271	TTT	AAA	CCC	AAT	TGC	AGG	GTC
272	Phe	Lys	Pro	Asn	Cys	Arg	Val
273			115			120	125
275	ACC	TGC	CTG	CCG	CTG	TGC	AGT
276	Thr	Cys	Leu	Pro	Leu	Cys	Ser
277		130				135	140
279	TGC	CCA	CGC	CCC	AAG	AGA	ATA
280	Cys	Pro	Arg	Pro	Lys	Arg	Ile
281	145				150		155
283	TGG	GTA	TGT	GAC	CAG	GGA	GTG

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/010,408

DATE: 02/14/2002

TIME: 19:09:10

Input Set : N:\Crif3\RULE60\10010408.txt

Output Set: N:\CRF3\02142002\J010408.raw

```

284 Trp Val Cys Asp Gln Gly Val Thr Pro Ala Ile Gln Arg Ser Thr Ala
285          165          170          175
287 CAA GGA CAC CAA CTT TCT GCC CTT GTC ACT CCT GCC TCT GCT GAT GCT      576
288 Gln Gly His Gln Leu Ser Ala Leu Val Thr Pro Ala Ser Ala Asp Ala
289          180          185          190
291 CCT TGT CCA AAT TGG AGC ACA GCC TGG GGC CCC TGC TCA ACC ACC TGT      624
292 Pro Cys Pro Asn Trp Ser Thr Ala Trp Gly Pro Cys Ser Thr Thr Cys
293          195          200          205
295 GGG CTG GGC ATA GCC ACC CGA GTG TCC AAC CAG AAC CGA TTC TGC CAA      672
296 Gly Leu Gly Ile Ala Thr Arg Val Ser Asn Gln Asn Arg Phe Cys Gln
297          210          215          220
299 CTG GAG ATC CAA CGC CGC CTG TGT CTG CCC AGA CCC TGC CTG GCA GCC      720
300 Leu Glu Ile Gln Arg Arg Leu Cys Leu Pro Arg Pro Cys Leu Ala Ala
301 225          230          235          240
303 AGG AGC CAC AGC TCA TGG AAC AGT GCT TTC TAA      753
304 Arg Ser His Ser Ser Trp Asn Ser Ala Phe
305          245          250
308 (2) INFORMATION FOR SEQ ID NO: 4:
310   (i) SEQUENCE CHARACTERISTICS:
311       (A) LENGTH: 8 amino acids
312       (B) TYPE: amino acid
313       (D) TOPOLOGY: linear
315   (ii) MOLECULE TYPE: peptide
317   (v) FRAGMENT TYPE: internal
321   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
W--> 323   Gly Cys Gly Cys Cys Xaa Xaa Cys
324       1          5
326 (2) INFORMATION FOR SEQ ID NO: 5:
328   (i) SEQUENCE CHARACTERISTICS:
329       (A) LENGTH: 177 base pairs
330       (B) TYPE: nucleic acid
331       (C) STRANDEDNESS: single
332       (D) TOPOLOGY: linear
334   (ii) MOLECULE TYPE: cDNA
337   (ix) FEATURE:
338       (A) NAME/KEY: CDS
339       (B) LOCATION: 1..177
342   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
344 TGT GAG GTG AAT GGC CGC AGG TAC CTG GAT GGA GAG ACC TTT AAA CCC      48
345 Cys Glu Val Asn Gly Arg Arg Tyr Leu Asp Gly Glu Thr Phe Lys Pro
346   1          5          10          15
348 AAT TGC AGG GTC CTG TGC CGC TGT GAT GAC GGT GGC TTC ACC TGC CTG      96
349 Asn Cys Arg Val Leu Cys Arg Cys Asp Asp Gly Gly Phe Thr Cys Leu
350          20          25          30
352 CCG CTG TGC AGT GAG GAT GTG CGG CTG CCC AGC TGG GAC TGC CCA CGC      144
353 Pro Leu Cys Ser Glu Asp Val Arg Leu Pro Ser Trp Asp Cys Pro Arg
354          35          40          45
356 CCC AAG AGA ATA CAG GTG CCA GGA AAG TGC TGC      177
357 Pro Lys Arg Ile Gln Val Pro Gly Lys Cys Cys

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/010,408

DATE: 02/14/2002

TIME: 19:09:11

Input Set : N:\Crf3\RULE60\10010408.txt

Output Set: N:\CRF3\02142002\J010408.raw

L:28 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:29 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]

L:323 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4

L:399 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7